## REMARKS

The present Amendment amends claims 3, 6 and 9. Therefore, the present application has pending claims 3, 6, and 9.

Claims 3, 6 and 9 stand rejected under 35 USC §101, as being directed to non-statutory subject matter. Amendments were made to claims 3, 6 and 9 to more clearly recite that each of the claims are directed to a statutory class of invention, namely a process (claim 3), machine (claim 6) and an article of manufacture (claim 9).

Further, amendments were made to claims 3, 6 and 9 so as to more clearly recite the practical application of the present invention as recited in the claims, wherein the invention is intended to find and retrieve a document similar to a seeds document by calculating a degree of similarity between the seeds document and each of the documents in the database and outputting a result including an indication of the degree of similarity of each of the documents of the database and the seeds document to thereby permit documents having a high degree of similarity to the seeds document to be retrieved.

Thus, claims 3, 6, and 9 are each directed to a permissible class of subject matter under 35 USC §101. Therefore, Applicants submit that this rejection with respect to claims 3, 6 and 9 is overcome and should be withdrawn

Claims 3, 6 and 9 stand rejected under 35 USC §112, second paragraph as being indefinite for failing to particularly point out and distinctly claim the subject matter which Applicants regards as their invention. Amendments were made claims 3, 6 and 9 to bring them into conformity with

8

the requirements of 35 USC §112, second paragraph. Therefore, Applicants submit that this rejection with respect to claims 3, 6 and 9 is overcome and should be withdrawn.

The Examiner's cooperation is respectfully requested to contact Applicants' Attorney by telephone should any further indefinite matters be discovered so that appropriate amendments may be made.

Claims 3, 6 and 9 stand rejected under 35 USC §102(b) as being anticipated by Noguchi (U.S. Patent No. 5,991,755). This rejection is traversed for the following reasons. Applicants submit that the features of the present invention as now more clearly recited in claims 3, 6 and 9 are not taught or suggested by Noguchi whether taken individually or in combination with any of the other references of record. Therefore, Applicants respectfully request the Examiner to reconsider and withdraw this rejection with respect to claims 3, 6 and 9.

Amendments were made to claims 3, 6, and 9 to clarify that the present invention is directed to a similar document search method, a similar document search system and a program product for making a computer operate as a similar document search system not taught or suggested by any of the references of record whether taken individually or in combination with each other.

Particularly, these claims now more clearly recite that the method, system and computer program implements, in a document search and retrieval system including a memory for storing a document database having documents and information concerning the documents and a processor for

processing data of the document database, searching of a document similar to a seeds document from the document database.

The present invention includes first extracting at least one characteristic word from the seeds document including desired retrieval contents; if the at least one characteristic word extracted by the first extracting step is one of a compound characteristic phrase constructed by a plurality of constituent characteristic words included in the compound characteristic phrase, second extracting of characteristic words included in said compound characteristic phrase and each of said plurality of constituent characteristic words of said compound characteristic phrase; calculating, according to the characteristic words extracted by the second extracting step, a degree of similarity between the seeds document and each of the documents stored on the database, by generating characteristics vectors using the characteristic words including the compound characteristic phrase the constituent characteristic words by which the compound characteristic phrase is constructed, wherein the characteristic vectors are extracted from the seeds documents and each of the documents stored on the document database and outputting a retrieval result, including an indication of the degree of similarity between each of the documents stored on the database and the seeds document, as a result of the degree of similarity between the seeds documents and each of the documents stored on the document database as calculated by the similarity calculating step, thereby permitting documents having a high degree of similarity to the seeds documents to be retrieved.

Further, according to the present invention the similarity calculating step includes the steps of calculating a weighting coefficient corresponding to

a distance between said constituent characteristic words on the document stored on the document database, said distance being calculated based on term appearance positions of said constituent characteristic words in a phrase including said characteristic words extracted in said second extracting step on the document stored on the document database; and calculating similarity using cosine of angles between the characteristic vectors extracted from the seeds documents and the characteristic vectors extracted from the document stored on the document database involving the weighting coefficient.

Thus, as per the features of the present invention as described above, first attention is directed to a compound characteristic phrase of the document. The compound characteristic phrase is constructed by a plurality of constituent characteristic words. If a characteristic word extracted by a first extracting step is one of the words of the compound characteristic phrase, then a second extracting step is conducted to extract characteristic words included in the compound characteristic phrase and each of the constituent characteristic words of the compound characteristic phrase. Further, as per the features of the present invention as described above, the distance used for calculations of weighting coefficients represents a distance between the constituent characteristic words of the compound characteristic phrase in the document to be retrieved.

The above described amendments to the claims so as to clarify the description of the present invention particularly with regard to the characteristic vectors can be found, for example, in Fig. 18 as elements 2300, 2031 and 2400 and the corresponding portions of the specification particularly, page 44, line 8 through page 47, line 2. Further, the description

of the amendments to the claims regarding using the cosine of an angle
between the characteristic vectors extracted from the seeds document and
the characteristic vectors extracted from each of the documents on the
database is described, for example, on page 47, line 1 through page 49, line
15 and as illustrated in Fig. 19.

The above described features of the present invention are not taught or
suggested by any of the references of record whether taken individually or in
combination with each other.  Particularly, the above described features of the
present invention are not taught or suggested by Noguchi.

Noguchi teaches, for example, at col. 12, lines 31-33 that a weighting
factor is set based on the occurrence frequencies of retrieval words.  In
Noguchi, the distance represents a distance between a characteristic vector
of seeds document and a characteristic vector of a document to be retrieved
in a characteristic vector space. Thus, there is no teaching or suggestion in
Noguchi that the distance being used is a distance between the constituent
characteristic words of the compound characteristic phrase in the document to
be retrieved and that this distance is used for calculations of weighting
coefficients corresponding to the distance.

Therefore, Noguchi clearly does not teach or suggest that a calculation
is performed to generate weighting coefficients corresponding to the distance
between the constituent characteristic words of the compound characteristic
phrase in the document to be retrieved as in the present invention as recited
in the claims.  Further, at no point in Noguchi is there any teaching or
suggestion that the weighting factor taught therein corresponds to a distance
calculated between term appearance positions on the seeds document as is

clearly recited in the claims. The weighting factors taught by Noguchi are simply intended to provide weighting of the frequency of occurrence of a retrieval word in the document.

Thus, similarity as calculated according to the present invention as recited in the claims is quite different from that of Noguchi in that in the present invention the distances between the terms is used to develop weighting coefficients so that, for example, terms having positions within the document that are closer together may have a higher weighting coefficient relative to terms having positions within the document that are farther apart. Such is clearly not taught or suggested by Noguchi since the weighting factors in Noguchi are simply to provide a higher weighting factor to terms that appear more often relative to terms that appear less often.

Further, in Noguchi, the similarity is a function of the functional unit collating unit 23. In the functional unit collating unit 23 of Noguchi, the similarity is judged only by checking whether or not a document includes two functional units. However, in the present invention, the similarity is calculated by using cosine of angle between the characteristic vectors extracted from the seeds document and the characteristic vectors extracted from the document stored on the document database.

Thus, Noguchi fails to teach or suggest calculating according to the characteristic words extracted by the second extracting step, a degree of similarity between the seeds document and each of the documents stored on the database, by generating characteristic vectors using the characteristic words including the compound characteristic phrase and the constituent characteristic words by which the compound characteristic phrase is

constructed, the characteristic vectors being extracted from the seeds document and each of the documents stored on the database and outputting a retrieval result, including an indication of the degree of similarity between each of the documents stored ont eh document database and the seeds, as a result of the degree of similarity between the seeds document and each of the documents stored on the database as calculated by the similarity calculating step, thereby permitting documents having a higher degree of similarity to the seeds document to be retrieved as recited in the claims.

Further, Noguchi fails to teach or suggest that if the at least one characteristic word extracted by the first extracting step is one of a compound characteristic phrase constructed by a plurality of constituent characteristic words included in the compound characteristic phrase, second extracting of characteristic words included in said compound characteristic phrase and each of said plurality of constituent characteristic words of said compound characteristic phrase as recited in the claims.

Still further, Noguchi fails to teach or suggest calculating a weighting coefficient corresponding to a distance between said constituent characteristic words on the document stored on the document database, said distance being calculated based on term appearance positions of said constituent characteristic words in a phrase including said characteristic words extracted in said second extracting step on the document stored on the document database as recited in the claims.

Even further yet, Noguchi fails to teach or suggest calculating similarity using cosine of angles between the characteristic vectors extracted from the seeds document and the characteristic vectors extracted from each of the

14

<u>document stored on the document database involving the weighting</u>

<u>coefficient</u> as recited in the claims.

Therefore, as is quite clear from the above, the features of the present invention as now more clearly recited in the claims are not taught or suggested by Noguchi whether taken individually or in combination with any of the other references of record and as such does not anticipate nor render obvious the features of the present invention as recited in the claims. Accordingly, reconsideration and withdrawal of the 35 USC §102(b) rejection of claims 3, 6 and 9 as being anticipated by Noguchi is respectfully requested.

The remaining references of record have been studied. Applicants submit that they do not supply any of the deficiencies noted above with respect to the reference utilized in the rejection of claims 3, 6 and 9.

In view of the foregoing amendments and remarks, applicants submit that claims 3, 6, and 9 are in condition for allowance. Accordingly, early allowance of claims 3, 6 and 9 is respectfully requested.

To the extent necessary, the applicants petition for an extension of time under 37 CFR 1.136. Please charge any shortage in fees due in connection with the filing of this paper, including extension of time fees, or credit any overpayment of fees, to the deposit account of MATTINGLY, STANGER, MALUR & BRUNDIDGE, P.C., Deposit Account No. 50-1417 (500.41226X00).

Respectfully submitted,

MATTINGLY, STANGER, MALUR & RUNDIDGE, P.C.

Carl I. Brundidge
Registration No. 29,621

CIB/jdc
(703) 684-1120

15